

COLLOCATION PROBLEMS IN POLICY CRITIQUES: A CORPUS-BASED STUDY

Bruce K C Ma, City University of Hong Kong
Roger C W Lee, Hong Kong Technical College

One of the causes of unnatural expressions found in tertiary students' writings is inappropriate lexical collocation. This lack of lexical competency is particularly significant in the work produced by learners of English as a second language. To rectify the problem, a systematic analysis of the problematic collocations is needed. This paper presents a project which looks at a 100,000-word corpus of policy critique assignments written by higher diploma students in Public Administration and Management at the City University of Hong Kong. A list of the most frequently used contents words is compiled with the help of the Longman Mini Concordancer. The suspected collocation problems associated with these words are put into a questionnaire. Eight native English speaking informants are invited to comment on the validity of the problematic collocations. The primary aim of this investigation is to produce a meaningful description of the nature and characteristics of the typical collocation problems in order to inform ESP course designers and teachers for curriculum revision. Analysis reveals that most collocation problems (1) are verb+noun structures, (2) fall into the category of CA (creation and/or activation) collocations, suggested by Benson (1990), and (3) are genre-specific, rather than topic-specific. These observations will be useful for designing teaching/learning materials to improve students' collocational competence in the specific genre.

Background

One of the obstacles to students learning English as a second language at the tertiary level is the lack of competence in putting words together properly. The problem is often reflected in the unnatural expressions which contain the right vocabulary items but in an inappropriate combination. For example, students may write 'accelerate ... speed' when they mean 'increase speed'. These inappropriate expressions give rise to the "foreign flavour" and are "the strongest markers of 'an accent'" (Korosadowicz-Struzynska (1980:115), quoted in Bahns (1993). Such problematic collocations have not been fully documented in the literature. A study carried out by Ma and Drave (1993) looked into the collocation problems of higher diploma students in Building Surveying and Architectural Studies at City University of Hong Kong. Useful insights were gained from this study and further investigation into learner writings of other disciplines was called for. The present project looks into the problem of collocation as evident in policy critique assignments produced by higher diploma students taking the Public Administration and Management Course.

Definition

Collocation is an intuitive ability of competent English speakers to put words in correct combinations. Such "an arbitrary and recurrent word combination" (Benson 1990) proves to be an obstacle to most ESL learners who are already too busy with the other rules of the English language when they use the language.

There are two types of collocations, namely grammatical and lexical (Benson 1985a, 1985b). Grammatical collocations are those which consist of a "dominant word" such as a verb, a noun or an adjective followed by a grammatical word, in most cases a preposition, as can be seen in examples like *sought after*, *deal with*, *take care of*, and so on. As to lexical collocations, they are usually formed by two "equal" lexical components in combinations like adjective+noun, noun+verb and verb+noun (Benson 1985a:62). Lexical collocations like *hold a meeting*, *conduct an interview*, *take the minutes*, and so on are the focus of the present research.

Research question

The immediate aim of this project is to use a learner corpus of around 100,000 words to identify problematic collocations in the writing of policy critiques, which are typical in Public Administration and Management

courses and careers. The findings will enable the investigators to establish a profile of lexical items which are most difficult for students to use in combination properly. This study will have an impact on modifying the teaching and learning of such courses so that both teachers and students are better aware of the specific collocation problems students have when writing in the specific genre of policy critique. If possible, the specific collocation problems will form the basis of materials development and design, which can be integrated with the original materials of the course to raise the learner's awareness of word combinability and increase their collocational competence.

Methodology

A 100,000 word corpus of learner writing in the specialised area of policy critique was assembled. This corpus was analysed with the Longman Mini-Concordancer (LMC), a computerised text analysis tool. Since LMC could only handle 50,000 words at a time, the corpus was divided into two parts and analysis was done separately and then results combined.

The corpus was created from policy critique assignments produced by Cantonese-speaking students studying the subject English Communication Skills in the course Higher Diploma in Public Administration and Management at City University of Hong Kong. Assignments were first collected in the form of electronic text or word-processed files. These assignments texts were then converted into LMC-readable ASCII text files. With the frequency word list generated by LMC, a list of the most frequently used content words in the genre of policy critique was identified. Concordance lines of these content words were then generated, from which suspected collocation problems were located. The problem statuses of these collocates were confirmed with the help of

- 1 Eight native speaker informants, who were asked to study the collocates and decide whether each one is acceptable or not; and
- 2 Four reference works, including
The BBI Combinatory Dictionary (Benson, Benson and Ilson, 1995)
A dictionary of English Collocations (Huang, 1991)
Longman Dictionary of English Collocations (Benson, Benson and Ilson, 1986)
The Word Finder (Rodale, 1947)

Once the problems were identified, attempts were made to produce a description of the problems, from which implications on course design, with particular reference to the teaching of the writing of the genre of policy critique, could be drawn. If possible, this description or typology will form the basis of supplementary work for learners, in the form of an independent learning package or materials to be integrated with original course materials, which aims at helping students become aware of the problems and mastering the correct collocates.

Findings

The LMC-generated frequency wordlist range from a top of 4,428 occurrences with the word "the" to a bottom of 1 with the word "youth's". Of this list, the most frequently used content words¹ were identified. They were: government (747), people (412), elderly (409), scheme (408), public (317), social (266), women (244), problem (229), housing (223), policy (215), services (183), many (167), labour (166), rate (157), workers (142), time (137), welfare (134), cent (132), society (129), pollution (125), education (124), increase (124), living (121), out (119), population (116), need (113), high (98), control (97), low (94), year (94), money (90), income (86), rent (86), class (84), market (82), pay (80), private (78), pension (74), property (70), work (69). The least frequently used word in this list occurs more than 60 times in the corpus. They are arguably words that learners have to use in writing this special genre of policy critique.

The concordance lines of the high frequency words were printed out for closer scrutiny. Suspected problematic collocations were identified on the basis of the researchers' intuition and experience with ESL students. These suspected collocation problems were then put together into a questionnaire for the native-speaking informants'. Altogether 78 concordance lines which covered 25 high frequency words were included.

Questionnaires (Appendix A) were given to eight informants, four from English language teaching background, and four from the Public Administration and Management disciplines. Each informant was given a souvenir as a token of their contribution.

¹ Definition of content words

Based on the native-speaking informants' feedback, 16 suspected problematic collocations were identified. The cut-off line for inclusion was set at "6X, 2?" meaning that six out of eight informants agreed that a particular concordance line carry a collocation mistake (indicated with "X") and two other informants were not sure (indicated with '?').

A summary of the problematic collocations is presented below.

Item	Status	Suspected problematic collocation
1	7X, 1?	<i>operate ... problem</i>
2	8X	<i>cheat ... money</i>
3	8X	<i>work popularly</i>
4	8X	<i>low-educated</i>
5	7X, 1V	<i>high intention</i>
6	7X, 1V	<i>pay ... care</i>
7	7X, 1?	<i>environment motivation</i>
8	7X, 1?	<i>government ... pleasurable</i>
9	7X, 1V	<i>taking abortion</i>
10	6X, 2?	<i>Give ... to devote</i>
11	7X, 1V	<i>demerits ... solve</i>
12	7X, 1?	<i>Get ... achievement</i>
13	7X, 1V	<i>accelerate ... speed</i>
14	8X	<i>deep review</i>
15	8X	<i>extremely rised</i>
16	7X, 1V	<i>enforce ... pressure</i>

Discussion

These problematic collocations reveal that most problems (1) are verb+noun structures, (2) fall into the category of CA(creation and/or activation) collocations, suggested by Benson (1990), and (3) are genre-specific. The following table gives an overview of the analysis. (See next page for table)

Eight out of the sixteen problems, i.e., 50 per cent of them, fall into the category of verb+noun collocations. The other kind of structure that attracts most problems is the adj+noun structure, which accounts for 3 occurrences.

Benson (1990) proposes a sub-classification under verb-noun collocations. CA collocations are those denoting creation, like *compile a dictionary*, *make an impression*, and those denoting activation, like *fly a kite*, *launch a missile*. EN collocations are those denoting eradication and/or nullification, like *reject an appeal*, *break a code*, *dispel fear*. In line with Benson's idea that CA collocations are "arbitrary and non-predictable" (Benson, 1985b: 64), students had a lot of problems with CA collocations. Five out of the eight verb-noun collocations fall into this category, while one belongs to the EN category.

If we assume that a primitive discursal analysis of the genre of policy critique could comprise three basic moves: (1) statements of the policy, (2) comments and criticisms on the policy and (3) suggestions and recommendations for the policy-makers, we could then conclude that there is a considerable degree of genre-specificity exhibited in the problems identified.

First, a significant number of collocation problems, five out of sixteen, relate to making suggestions and recommendations. These include: *to operate a problem*, *to solve demerits*, *to accelerate speed*, *to give the government to do sth*, and *to enforce pressure*.

Another important insight we can draw from this study is the abundance of collocation problems relating to "intensity words". Policy critiques typically involving making comments on existing policies. This means saying whether a certain policy is good or bad, successful or ineffective. Inevitably words expressing intensity have to be used. Learners apparently have in their vocabulary a wide range of words denoting different levels of intensity, like *large/small*, *high/low*, *good/bad*, *heavy/light*, but to use these intensity words correctly with

the object or idea they are talking about is certainly a very difficult task, as evident in the number of problems falling into this category. There are incidentally many other suspected collocation problems in this area, though they have not occurred with the most frequently used content words to warrant inclusion in our analysis. The following are some examples: *low amount of assistance, low-skill jobs, low punishment, low subsidy, low ability, high penalties, high popularity, a fuller life, and highly violated.*

Unlike the results of a study by Ma (1997), where a corpus of site investigation reports produced by students studying Building-related disciplines revealed a multitude of topic-specific collocation problems, problems identified in this study bear very little topic-specificity. Only two of the eighteen problems could be grouped under a specific topic that learners were writing about: *the environmental motivation* from the topic of conservation and *to take abortion* from the topic of abortion.

Collocation	Analysis	Grammatical	CA/EN	Genre	Topic
1. <i>to operate a problem</i>		Verb-noun		Suggestions	
2. <i>to work popularly</i>		Verb-adverb			
3. <i>a high intentions</i>		Adj-noun		Intensity	
4. <i>the environmental motivation</i>		Adj-noun			Conservation
5. <i>to take abortion</i>		Verb-noun	CA		Abortion
6. <i>to solve demerits</i>		Verb-noun	EN	Suggestions	
7. <i>to accelerate speed</i>		Verb-noun	CA	Suggestions	
8. <i>prices to rise extremely</i>		Verb-adverb		Intensity	
9. <i>to cheat money</i>		Verb-noun			
10. <i>to be low-educated</i>		Adverb-adj		Intensity	
11. <i>to pay care</i>		Verb-noun	CA		
12. <i>the government being pleasurable</i>		Noun-adj			
13. <i>to give the government to do sth</i>		Verb-verb		Suggestions	
14. <i>to get achievement</i>		Verb-noun	CA		
15. <i>a deep review</i>		Adj-noun		Intensity	
16. <i>to enforce pressure</i>		Verb-noun	CA	Suggestions	

Conclusion

From the previous analysis and discussion, we can conclude that collocation problems learners in this study have exhibit an identifiable pattern. There are plenty of problems with verb-noun collocations, CA collocations in particular, and the problems are mostly genre-specific rather than topic-specific. The fact that we are only examining a small corpus and a list of only 16 frequently used content words and still are able to summarize a pattern strengthens our belief that these problems could apply across a wider population of learners attempting to write policy critiques.

The large number of CA verb-noun collocation problems only substantiates Benson’s argument that they are “arbitrary and non-predictable” and are particularly difficult for ESL learners. This does not imply a comparison with other types of collocations like adj-+noun or EN collocations. The fact that many problems fall into this category could simply mean that there is the need to use this type of collocation more often than others in the genre of policy critique. Benson suggests that “a guide” be given to the learners to help them master these collocations and we believe the data from this study will shed some light on what collocation items should be included in this guide.

The lack of topic-specificity suggests that learners writing this particular genre of policy critique are not bothered by discipline-specific, or ESP, collocations as much as they are by general English, or EAP (English for Academic Purposes), collocations. One reason could be that learners in the study were very heavily exposed to readings from newspapers and magazines on the topic or policy they are critiquing about before attempting to write. In this way, discipline-specific and topic-specific collocations are well learnt in chunks before writing.

Of the collocation problems we identified, there are incidentally quite a number that we could attribute to first language influence, i.e. we believe learners use English equivalents of words they use in Cantonese where these words correctly collocate. These include: *to cheat money*, *to be low-educated*, *to give the government to do sth*, *to get achievement*, and *a deep review*.

To give assistance to learners learning to write policy critiques, it could be suggested that in the least a list of the correct collocations from the problems identified in this study be compiled and explained to students before they are assigned to write the critique. For better results, some self-access materials on collocations could be developed systematically highlighting the CA and genre-specific collocations learners will often need to use. Moreover, learners could be made aware of the influence of Cantonese collocation expressions which, when directly translated, do not collocate in English

In the light of the results of this study and Ma and Drave's study (1993), we believe there is the need to more systematically set up corpus databases for research that would include learner writings from a wider range of representative genres as well as from general English or EAP. Studies like the present one and Ma and Drave's (1993) are specific in the learner groups and target genre. They do inform ESP course designers and teachers about the learner problems and suggest directions for course materials addition and revision. However, the observation in this study that learners have more problems with general English collocations rather than ESP collocations, which runs contrary to Ma and Drave's study, has dispelled our intuition that ESP students need more help with ESP collocations than general English collocations. It is clear that more corpus-based work need to be done across a much wider spectrum of learner population in order to better define and describe learners' collocation problems and validate or invalidate the ESP/EAP dichotomy. Until then course designers could only take hint from studies like this one to include in their course presentation, in one form or another, of the correct collocations and attempt to increase learners' awareness of the incorrect ones.

References

- Bahns, J. & Eldaw, M. (1993). Should we teach EFL students collocations? *System*, 21(1), 101-114.
- Bahns, J. (1993). Lexical collocations: a contrastive view. *ELT Journal*, 47(1), 56-63.
- Benson, M. (1985a). Collocation and idioms. In R. Ilson (Ed.), *Dictionaries, lexicography and language learning*. Oxford: British Council and Pergamon.
- Benson, M. (1985b). Lexical combinability. In W. J. Frawley & R. Steiner, (Eds.), *Advances in lexicography*. Special issue of *Papers in Linguistics*, 18(1), 3-15.
- Benson, M. (1990). Collocations and general-purpose dictionaries. *International Journal of Lexicography*, 3(1), 23-35.
- Benson, M., Benson, E., & Ilson, R. (1986). *The BBI combinatory dictionary of English. A guide to word combinations*. Amsterdam/Philadelphia: Benjamins.
- Benson, M., Benson, E., & Ilson, R. (1995). *Longman dictionary of English collocations (English-Chinese)*. Hong Kong: Longman.
- Huang, W. (1991). *A dictionary of English collocations*. Jiangsu: Jiangsu Educational Press.
- Ma, B. (1997). Building "Building Words" – a computer-assisted learning package for ESP collocation. *English Language Studies Working Papers*, 2(1), Division of Language Studies, City University of Hong Kong.
- Ma, B. and Drave, N. (1993). Collocation problems in technical writing. Paper presented at Second International Conference on English for Professional Communication. City Polytechnic of Hong Kong, March, 1993.
- Rodale, J. I. (1947). *The word finder*. Emmaus, Pennsylvania: Rodale Books Inc.